

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|  |           |   |
|--|-----------|---|
| <b>(51) International Patent Classification <sup>5</sup> :</b><br><b>C12N 15/81, 1/16, 9/42</b>  | <b>A1</b> | <b>(11) International Publication Number:</b> <b>WO 93/11249</b><br><b>(43) International Publication Date:</b> 10 June 1993 (10.06.93)   |
| <b>(21) International Application Number:</b> PCT/DK92/00360<br><b>(22) International Filing Date:</b> 2 December 1992 (02.12.92)<br><br><b>(30) Priority data:</b><br>PCT/DK91/00379 4 December 1991 (04.12.91) WO<br>(34) Countries for which the regional or international application was filed: DK et al.<br>PCT/DK91/00378 4 December 1991 (04.12.91) WO<br>(34) Countries for which the regional or international application was filed: DK et al.<br><br><b>(71) Applicant (for all designated States except US):</b> NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsvaerd (DK).   |           | <b>(72) Inventors; and</b><br><b>(75) Inventors/Applicants (for US only) :</b> DALBØGE, Henrik [DK/DK]; Ligustervænget 63, DK-2830 Virum (DK). HELDT-HANSEN, Hans, Peter [DK/DK]; Vangedet 53, DK-2830 Virum (DK). RASMUSSEN, Grethe [DK/DK]; Theklavej 15, st.tv., DK-2400 København NV (DK).<br><br><b>(81) Designated States:</b> BR, CA, FI, JP, KR, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).<br><br><b>Published</b><br><i>With international search report.</i><br><i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> |
| <b>(54) Title:</b> A METHOD OF CLONING PROTEINS IN YEAST AND CELLULASE FROM HUMICOLA INSOLENS<br><br><b>(57) Abstract</b><br><br>Method of screening for a DNA sequence coding for a protein of interest, the method comprising a) cloning, in suitable vectors, a DNA library from an organism suspected of producing one or more proteins of interest; b) transforming suitable yeast host cells with said vectors; c) culturing the host cells under suitable conditions to express any protein of interest encoded by a clone in the DNA library, and d) screening for positive clones by determining any activity of a protein expressed in step (c). An enzyme which exhibits cellulase activity and has been isolated from DNA library of Humicola insolens. The enzyme has a cellulose binding domain and exhibits endocellulase activity in the presence of linear alkyl benzene sulfonate. |           |   |

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                       |    |                          |
|----|--------------------------|----|---------------------------------------|----|--------------------------|
| AT | Austria                  | FR | France                                | MR | Mauritania               |
| AU | Australia                | GA | Gabon                                 | MW | Malawi                   |
| BB | Barbados                 | GB | United Kingdom                        | NL | Netherlands              |
| BE | Belgium                  | GN | Guinea                                | NO | Norway                   |
| BF | Burkina Faso             | GR | Greece                                | NZ | New Zealand              |
| BG | Bulgaria                 | HU | Hungary                               | PL | Poland                   |
| BJ | Benin                    | IE | Ireland                               | PT | Portugal                 |
| BR | Brazil                   | IT | Italy                                 | RO | Romania                  |
| CA | Canada                   | JP | Japan                                 | RU | Russian Federation       |
| CF | Central African Republic | KP | Democratic People's Republic of Korea | SD | Sudan                    |
| CG | Congo                    | KR | Republic of Korea                     | SE | Sweden                   |
| CH | Switzerland              | KZ | Kazakhstan                            | SK | Slovak Republic          |
| CI | Côte d'Ivoire            | LI | Liechtenstein                         | SN | Senegal                  |
| CM | Cameroon                 | LK | Sri Lanka                             | SU | Soviet Union             |
| CS | Czechoslovakia           | LU | Luxembourg                            | TD | Chad                     |
| CZ | Czech Republic           | MC | Monaco                                | TG | Togo                     |
| DE | Germany                  | MG | Madagascar                            | UA | Ukraine                  |
| DK | Denmark                  | ML | Mali                                  | US | United States of America |
| ES | Spain                    | MN | Mongolia                              | VN | Viet Nam                 |
| FI | Finland                  |    |                                       |    |                          |

**BEST AVAILABLE COPY**

A. method of cloning proteins in yeast and cellulase from *Humicola insolens*.

#### FIELD OF INVENTION

- 5 The present invention relates to a method of screening for DNA sequences coding for proteins of interest, as well as to a process for producing such proteins of interest.

#### BACKGROUND OF THE INVENTION

10

The advent of recombinant DNA techniques has made it possible to select single protein components with interesting properties and produce them on a large scale. This represents an improvement over the previously employed production process  
15 using microorganisms isolated from nature and producing a mixture of proteins which would either be used as such or separated after the production step. However, the conventional cloning techniques have the drawback that each protein component has to be purified and characterized by its (partial)  
20 amino acid sequence before it is possible to prepare synthetic oligonucleotide probes for hybridization experiments. Since this is a rather time-consuming process, the cloning of novel proteins might be considerably expedited by using a screening method involving selecting clones expressing a desired protein  
25 activity.

Such a screening method has previously been devised for the cloning of prokaryotic gene products in Bacillus, cf. US 4,469,791; P. Cornelis et al., Mol. Gen. Genet. 186, 1982, pp. 507-511; I. Palva, Gene 19, 1982, pp. 81-87; S.A. Ortlepp, Gene 23, 1983, pp. 267-276; H. Yamazaki et al., J. Bacteriol. 156, 1983, pp. 327-337; N. Tsukagoshi et al., Mol. Gen. Genet. 193, 1984, pp. 58-63; M. Sibakov and I. Palva, Eur. J. Biochem. 145, 1984, pp. 567-572; and J.R. Mielenz, Proc. Natl. Acad. Sci. USA  
35 80, 1983, pp. 5975-5979. A screening method based on expression cloning of eukaryotic genes in mammalian cells has been described, e.g. in D.P. Gearing et al., The EMBO J. 8, 1989,

pp. 3667-3676; N. Harada et al., Proc. Natl. Acad. Sci. USA 87, 1990, pp. 857-861; and R. Fukunaga et al., Cell 61, 1990, pp. 341-350.

## 5 SUMMARY OF THE INVENTION

It has now been found possible to screen for yeast clones expressing protein activities of interest with a view to isolating DNA coding for single protein components.

10

Accordingly, the present invention relates to a method of screening for a DNA sequence coding for a protein of interest, the method comprising

15 (a) cloning, in suitable vectors, a DNA library from an organism suspected of producing one or more proteins of interest,

(b) transforming suitable yeast host cells with said vectors,

20

(c) culturing the host cells under suitable conditions to express any protein of interest encoded by a clone in the DNA library, and

25 (d) screening for positive clones by determining any activity of a protein expressed in step (c).

As indicated above, expression cloning of prokaryotic genes in Bacillus has previously been described. The prokaryotic systems  
30 devised for expression cloning, however, are not operable for the cloning of eukaryotic genes which are generally difficult to express in Bacillus. While expression cloning of eukaryotic genes in mammalian cells has been described, it is more advantageous to use yeast as a host organism as it is possible  
35 to obtain a much higher transformation frequency than with mammalian cells, and as yeast is far easier to cultivate. Furthermore, the yeast clones are stable whereas the mammalian

expression cloning system described in the references cited above is based on transient expression in COS cells. Unlike the mammalian system, the yeast system results in pure clones after the initial screening and, therefore, they need not be screened  
5 in pools and subpools as in the mammalian system. Apart from this conventional selection systems may be used to select yeast transformants.

According to the present invention, it has surprisingly been  
10 found that yeast cells appear to be able to express heterologous genes extracellularly by means of heterologous secretion signals in amounts which are sufficient for screening purposes. Although expression cloning of certain proteins in yeast has been described previously (G.L. McKnight and B.L.  
15 McConaughy, Proc. Nat. Acad. Sci. USA 80, 1983, pp. 4412-4416), it has not been generally useful as it is based on complementation of essential genes and therefore is dependent on yeast host strains which have been mutated to lack these essential genes. In the present screening method, no such  
20 requirement is necessary for the yeast host strain to be used in the method. Besides, the gene products of the previously described method are intracellular rather than extracellular as in the present method.

25 The advantage presented by the present screening method is primarily that it requires no prior knowledge of the structure of the protein of interest. This means that the rate at which novel genes may be isolated and, consequently, novel products be developed may be greatly increased. Furthermore, the method  
30 permits screening for multiple protein activities and may even result in the isolation of several different genes coding for the same type of proteins.

In another aspect, the present invention relates to a process  
35 for producing a protein of interest in a heterologous host cell, the process comprising transforming a suitable heterologous host cell with a DNA sequence coding for a protein

of interest, which DNA sequence has been isolated by the screening method of the invention, culturing the transformed cells under suitable conditions to express the protein, and recovering the expressed protein from the culture.

5

In a further aspect, the present invention relates to an enzyme which exhibits cellulase activity, and which has the following characteristics

10 (a) the DNA sequence encoding the enzyme has been isolated from a DNA library of Humicola insolens,

(b) said DNA sequence comprises at least one of the following partial sequences

15

(i) TGGCAGCAGT GTGGTGGCGT TGGCTTCTCG GGCTCTACGT  
CCTGTGTGTC CGGTTACACG TCGGTGTACT TGAACGACTG  
GTACAGCCAA TGC  
(SEQ ID#1)

20

(ii) CAGCGCAGCC GACGACGTTA CGGACAACAC AACAAACGACC  
AGGGCAACAT CGACAACAAG GTCAGCCCCG GCTGCCACTT  
CAACCACTCCG G  
(SEQ ID#2)

25

(iii) CCAAGGCGAA GTTCAAGTGG TTGGCATCAA CCAGTCCTGC  
GCTGAGTTCG GCAAGGAGAG TATCCGGCTA TGGGCAAGCA  
CTTACTTCCT TCGCGACGTC GTCGATTCAA GCGCACATCA  
ATCGTGGCTT CA (SEQ ID#3)

30

(iv) CTGACGTGAA CGTGACCAAC AACAACTTGG CCGTAGCGAC  
CGAGAACAAG CTGTGTACCA GATGCATCA (SEQ ID#4)

35

(v) GGACGGTCCG GCACGAGCAC GGCCTGCGTC AGCACCCAGG  
TCGGCCTTCA GCGCGTCATT GCGCGACCA ACTGGCTCAG  
GCAAAACGGC AAGGTTGGAC TGCTCGCGAC TTGCCGCGGC (SEQ  
ID#5)

(vi)       GCCAAGTGGG     TTTGCCAGCA     GGCCATTGAG     GGCATGCTGA  
          ACCACCTCCA     GGAGAATAGC     GATGTCTGGA     CAGGTGCGCT  
          CTGGTGGGCG     GGAGGCCCGT     GGTGGGGTTG     ACTATATCTA (SEQ  
          ID#6)

5

(c) the enzyme comprises a cellulose-binding domain, and

(d) the enzyme exhibits endocellulase activity in the presence of linear alkyl benzene sulfonate.

10

The enzyme of the invention may be isolated by the method of the invention.

In the present context, the term "cellulose-binding domain" is intended to indicate an amino acid sequence capable of effecting binding of the enzyme to a cellulosic substrate. Cellulose-binding domains have been found to be important for the endoglucanase activity of cellulytic enzymes on substrates (cf. the discussion in PCT/DK91/00124). The term "endocellulase activity" refers to the ability of the enzyme to degrade cellulose to glucose, cellobiose, triose and other cello-oligosaccharides, as determined by the formation of clearing zones in a carboxymethyl cellulose (CMC) gel under the conditions specified below. Unlike the endocellulase described in PCT/DK91/00123), the enzyme of the present invention shows substantially unchanged stability in the presence of linear alkyl benzene sulfonates. This is an important advantage as linear alkyl benzene sulfonates are commonly used in detergent compositions.

30

#### DETAILED DISCLOSURE OF THE INVENTION

According to the invention, the DNA library is preferably a cDNA library prepared from the mRNA of an organism suspected of producing one or more proteins of interest. Although it may also be possible to screen genomic libraries in this manner, at least some potential yeast hosts may not be able to splice

35

eukaryotic genomic DNA correctly, and therefore a positive result of the screening may more often be obtained by using cDNA instead.

- 5 To ensure a more accurate result, it may be an advantage to subject positive clones isolated in step (d) of the present method to rescreening, reisolation and recloning.

The organism suspected of producing one or more proteins of  
10 interest is typically a eukaryotic organism, in particular a fungus since fungi are known to produce a large number of different proteins which makes the traditional process of isolating a gene coding for a particular protein product by initially purifying each protein separately particularly  
15 cumbersome. This makes it particularly advantageous to screen fungal DNA libraries by the method of the invention because a large number of different protein activities (and DNAs coding for them) may be identified within a relatively short time-span using the same library. In this respect, screening of yeast  
20 colonies for different protein activities is far more efficient than screening of filamentous fungi as a large number (i.e. about 500-1000) of yeast colonies may be grown on each plate, compared to 10-50 filamentous fungi/plate.

25 One type of industrially useful proteins currently obtained from fungi is enzymes. Thus, yeast clones may be screened by the method of the invention for expression of one or more enzyme activities by means of appropriate assays. Examples of enzymes which may be identified by this method are  
30 carbohydrases, e.g. cellulytic enzymes such as endocellulases, cellobiohydrolases,  $\beta$ -glucanases or  $\beta$ -glucosidases, hemicellulases or pectinolytic enzymes such as galactanases, galactosidases, mannanases, xylanases, pectinases, xylosidases, arabanases, rhamnogalacturonases or amylases; esterases, e.g.  
35 lipolytic enzymes such as lipases; proteases; oxidoreductases, e.g. peroxidases, oxidases or laccases; or isomerases, e.g. glucose isomerase.



A wide range of indicator systems for the different types of enzymes may be used for the screening of yeast colonies on agar plates. For instance, endocellulases may be identified by clearing zones in carboxymethyl cellulose after staining with Congo Red; similar methods may be used to detect glucanases, xylanases and galactanases. Endoarabanases may be identified by blue zones obtained after dissolution of azurine-crosslinked araban. This principle is general and may be used to detect, e.g., mannanases, xylanases and cellulases. Pectinases (polygalacturonases and pectin lyases) may be identified by clearing zones in pectin after precipitation with quaternary ammonium ions. Amylases may be identified by clearing zones in starch after visualisation with iodine.  $\alpha$ -galactosidases may be detected by the release of p-nitrophenol (yellow) from p-nitrophenol- $\alpha$ -galactopyranoside or by coupling released naphthole or naphthole derivatives from, e.g., 1-naphthole- $\alpha$ -galactopyranoside to azo dyes; similar methods may be used to detect  $\beta$ -galactosidases,  $\alpha$ - and  $\beta$ -glycosidases,  $\beta$ -xylosidase and  $\beta$ -mannosidase. Numerous methods are available for the detection of proteases, e.g. clearing zones in casein after precipitation with trichloroacetic acid. Peroxidases and oxidases may be detected by the reaction of 4-aminoantipyrine with ESBT (N-ethyl-N-sulfoethyl-m-toluidine) in the presence of hydrogen peroxide (generating a purple colour). Lipases may be detected by the formation of clearing zones in tributyrine emulsions.

The yeast strain selected to be the host cell for the DNA library may be any yeast strain conventionally used for the cloning of heterologous DNA sequences. Thus, the yeast strain may suitably be selected from Saccharomyces sp., such as Saccharomyces cerevisiae, Saccharomyces kluyveri, Saccharomyces uvarum or Schizosaccharomyces pombe, Hansenula sp. Pichia sp., Yarrowia sp. such as Yarrowia lipolytica, or Kluyveromyces sp. such as Kluyveromyces lactis.

The vector used for cloning the DNA library may be any vector which may conveniently be subjected to recombinant DNA procedures. In each vector, the DNA sequence derived from the library should be operably connected to a suitable promoter  
5 sequence. The promoter may be any DNA sequence which shows transcriptional activity in the yeast cell. Examples of suitable promoters for use in yeast host cells include promoters from yeast glycolytic genes (Hitzeman et al., J. Biol. Chem. 255, 1980, pp. 12073-12080; Alber and Kawasaki, J. Mol. Appl. Gen. 1, 1982, pp. 419-434) or alcohol dehydrogenase  
10 genes (Young et al., in Genetic Engineering of Microorganisms for Chemicals (Hollaender et al, eds.), Plenum Press, New York, 1982), or the TPI1 (US 4, 599, 311) or ADH2-4c (Russell et al., Nature 304, 1983, pp. 652-654) promoters.

15 Each DNA library sequence may also be operably connected to a suitable terminator, such as the TPI1 (Alber and Kawasaki, op. cit.) or ADH3 (McKnight et al., op. cit.) or yeast MFA terminators.

20 The vector may further comprise a DNA sequence enabling the vector to replicate in yeast cell. An example of such a sequence is the yeast plasmid 2 $\mu$  replication genes REP 1-3 and origin of replication. If the vector is a yeast/E. coli shuttle  
25 vector, it will also include an origin of replication region which is functional in E. coli. The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell such as URA3, or one which confers resistance to a drug, e.g. ampicillin, kanamycin,  
30 chloramphenicol, tetracyclin, etc., or the Schizosaccharomyces pombe TPI gene (described by P.R. Russell, Gene 40, 1985, pp. 125-130).

The procedures used to ligate the DNA library sequences, the  
35 promoter and the terminator, respectively, and to introduce them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art

(cf., for instance, Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989). The transformation of yeast cells may for instance be effected by protoplast formation followed by transformation or by the LiAc method in a manner known per se.

In the process of the invention of producing a protein of interest after the DNA coding for the protein has been isolated by the screening method described above, the heterologous host cell transformed with the isolated DNA sequence may be a strain of a filamentous fungus, e.g. fungi belonging to the groups Phycomycetes, Zygomycetes, Ascomycetes, Basidiomycetes or Fungi Imperfecti, including Hyphomycetes such as the genera Aspergillus, Trichoderma, Penicillium, Fusarium or Humicola.

The filamentous fungus host organism may conveniently be one which has previously been used as a host for producing recombinant proteins, e.g. a strain of Aspergillus sp., such as A. niger, A. nidulans or A. oryzae. The use of A. oryzae in the production of recombinant proteins is extensively described in, e.g. EP 238 023.

In particular when the host organism is A. oryzae, a preferred promoter for use in the process of the present invention is the A. oryzae TAKA amylase promoter as it exhibits a strong transcriptional activity in A. oryzae. The sequence of the TAKA amylase promoter appears from EP 238 023.

Termination and polyadenylation sequences may suitably be derived from the same sources as the promoter.

The techniques used to transform a fungal host cell may suitably be as described in EP 238 023.

The medium used to culture the transformed host cells may be any conventional medium suitable for growing Aspergillus cells. The mature protein secreted from the host cells may

conveniently be recovered from the culture medium by well-known procedures including separating the cells from the medium by centrifugation or filtration, and precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, followed by chromatographic procedures such as ion exchange chromatography, affinity chromatography, or the like.

A preferred endocellulase enzyme according to the invention is an enzyme, a crude extract (15  $\mu$ l) of which diluted with one volume of 0.15% linear alkyl benzene sulfonate and added to a 2% agarose gel containing 1% carboxymethyl cellulose in 50 mM sodium phosphate buffer, pH 7, mixed with one volume of 0.15% linear alkyl sulfonate forms a clearing zone in said agarose gel after 18 hours of incubation, which clearing zone is equal to (less 3 mm) the clearing zone formed in a similar carboxymethyl cellulose gel not containing any linear alkyl benzene sulfonate, provided that the concentration of enzyme in the extract is such that a clearing zone of at least 10 mm is formed in a carboxymethyl cellulose gel (with no linear alkyl benzene sulfonate) under the conditions specified above.

The DNA sequence coding for the enzyme may for instance be isolated by screening a cDNA library of Humicola insolens, e.g. strain DSM 1800, deposited on 1 October 1981 at the Deutsche Sammlung von Mikroorganismen in accordance with the provisions of the Budapest Treaty and selecting for clones expressing the appropriate enzyme activity (i.e. endocellulase activity as defined above). The appropriate DNA sequence may then be isolated from the clone by standard procedures, e.g. as described in Example 1.

In a further aspect, the invention relates to a detergent additive comprising the enzyme of the invention. The detergent additive may suitably be in the form of a non-dusting granulate, stabilized liquid or protected enzyme. Non-dusting granulates may be produced e.g. according to US 4,106,991 and 4,661,452 (both to Novo Industri A/S) and may optionally be

coated by methods known in the art. Liquid enzyme preparations may, for instance, be stabilized by adding a polyol such as propylene glycol, a sugar or sugar alcohol, lactic acid or boric acid according to established methods. Other enzyme  
5 stabilizers are well known in the art. Protected enzymes may be prepared according to the method disclosed in EP 238 216.

It will be understood that the detergent additive may further include one or more other enzymes, such as a protease, lipase,  
10 peroxidase or amylase, conventionally included in detergent additives.

In a still further aspect, the present invention relates to a detergent composition comprising the enzyme of the invention.  
15 The detergent composition of the invention may be in any convenient form, e.g. as powder, granules or liquid. A liquid detergent may be aqueous, typically containing up to 90% water and 0-20% organic solvent.

20 The detergent composition comprises a surfactant which may be anionic, non-ionic, cationic, amphoteric or a mixture of these types. The detergent will usually contain 0-50% anionic surfactant such as linear alkyl benzene sulphonate (LAS), alpha-olefin sulphonate (AOS), alkyl sulphate (AS), alcohol  
25 ethoxy sulphate (AES) or soap. It may also contain 0-40% non-ionic surfactant such as nonyl phenol ethoxylate or alcohol ethoxylate. Furthermore, it may contain a polyhydroxy fatty acid amide surfactant (e.g. as described in WO 92/06154).

30 The detergent composition may additionally comprise one or more other enzymes, such as an amylase, lipase, peroxidase, oxidase or protease.

The pH (measured in aqueous detergent solution) will usually  
35 be neutral or alkaline, e.g. 7-11. The detergent may contain 1-40% of a detergent builder such as zeolite, phosphate, phosphonate, citrate, NTA, EDTA or DTPA, alkenyl succinic

anhydride, or silicate, or it may be unbuilt (i.e. essentially free from a detergent builder). It may also contain other conventional detergent ingredients, e.g. fabric conditioners, foam boosters, bleaching agents, e.g. perborate, percarbonate, 5 tetraacetyl ethylene diamine (TAED), or nonanoyloxybenzene sulfonate (NOBS), anti-corrosion agents, soil-suspending agents, sequestering agents, anti-soil redeposition agents, stabilizing agents for the enzyme(s), foam depressors, dyes, bactericides, optical brighteners or perfumes.

10

Particular forms of detergent composition within the scope of the invention include:

a) A detergent composition formulated as a detergent powder 15 containing phosphate builder, anionic surfactant, nonionic surfactant, silicate, alkali to adjust to desired pH in use, and neutral inorganic salt.

b) A detergent composition formulated as a detergent powder 20 containing zeolite builder, anionic surfactant, nonionic surfactant, acrylic or equivalent polymer, silicate, alkali to adjust to desired pH in use, and neutral inorganic salt.

c) A detergent composition formulated as an aqueous detergent 25 liquid comprising anionic surfactant, nonionic surfactant, humectant, organic acid, caustic alkali, with a pH in use adjusted to a value between 7 and 10.5.

d) A detergent composition formulated as a nonaqueous deter- 30 gent liquid comprising a liquid nonionic surfactant consisting essentially of linear alkoxylated primary alcohol, phosphate builder, caustic alkali, with a pH in use adjusted to a value between about 7 and 10.5.

e) A detergent composition formulated as a detergent powder 35 in the form of a granulate having a bulk density of at least 600 g/l, containing anionic surfactant and nonionic surfactant,

low or substantially zero neutral inorganic salt, phosphate builder, and sodium silicate.

- f) A detergent composition formulated as a detergent powder in the form of a granulate having a bulk density of at least 600 g/l, containing anionic surfactant and nonionic surfactant, low or substantially zero neutral inorganic salt, zeolite builder, and sodium silicate.
- g) A detergent composition formulated as a detergent powder containing anionic surfactant, nonionic surfactant, acrylic polymer, fatty acid soap, sodium carbonate, sodium sulphate, clay particles, and sodium silicate.
- h) A liquid compact detergent comprising 5-65% by weight of surfactant, 0-50% by weight of builder and 0-30% by weight of electrolyte.

Apart from these ingredients, the detergent compositions a)-h) include the cellulase of the invention and optionally one or more other enzymes, as indicated above.

The softening, soil removal and colour clarification effects obtainable by means of the enzyme of the invention generally require a concentration of the enzyme in the washing solution of 0.0001 - 100, preferably 0.0005 - 60, and most preferably 0.01 - 20 mg of enzyme protein per liter. The detergent composition of the invention is typically employed in concentrations of 0.5 - 20 g/l in the washing solution. In general, it is most convenient to add the detergent additive in amounts of 0.1 - 5% w/w or, preferably, in amounts of 0.2 - 2% of the detergent composition.

#### BRIEF DESCRIPTION OF THE DRAWINGS

35

Fig. 1 is a map of plasmid pYHD17, wherein "TPI promoter" indicates the S. cerevisiae triose phosphate isomerase

promoter, "Terminator" indicates the S. cerevisiae triose phosphate isomerase terminator, "Amp" indicates the gene mediating ampicillin resistance, "2 $\mu$  ori" indicates the yeast plasmid 2 $\mu$  origin of replication, and "URA3" indicates a gene  
5 encoding a selection marker complementing a uracil deficiency in the host strain; and

Fig. 2 is a map of plasmid pHD414, wherein "AMG Terminator" indicates the A. niger glucoamylase terminator, and "TAKA  
10 Promoter" indicates the A. oryzae TAKA amylase promoter;

The present invention is further illustrated in the following examples which are not in any way intended to limit the scope of the invention as claimed.

15

#### EXAMPLES

##### **Materials and Methods**

20 Donor organism: mRNA was isolated from the following organisms: H. insolens, DSM 1800, grown in a cellulose-rich fermentation medium with agitation to ensure sufficient aeration.

Construction of an expression plasmid: The commercially  
25 available plasmid pYES II (Invitrogen) was cut with SpeI, filled in with Klenow DNA polymerase + dNTP and cut with ClaI. The DNA was size fractionated on an agarose gel, and a fragment of about 2000 bp was purified by electroelution. The same plasmid was cut with ClaI/PvuII, and a fragment of about 3400  
30 bp was purified by electroelution. The two fragments were ligated to a blunt-ended SphI/EcoRI fragment containing the yeast TPI promoter. This fragment was isolated from a plasmid in which the TPI promoter from S. cerevisiae (cf. T. Albers and G. Kawasaki, J. Mol. Appl. Genet. 1, 1982, pp. 419-434) was  
35 slightly modified: an internal SphI site was removed by deleting the four bp constituting the core of this site. Furthermore, redundant sequences upstream of the promoter were



removed by Ball exonuclease treatment followed by addition of a SphI linker. Finally, an EcoRI linker was added at position -10. After these modifications, the promoter is included in a SphI-EcoRI fragment. Its efficiency compared to the original promoter appears to be unaffected by the modifications. The resulting plasmid pYHD17 is shown in Fig. 1.

**Isolation of mRNA:** Total RNA was isolated from approximately 7 g of mycelium. The mycelium was frozen in liquid nitrogen and ground in a mortar with 1 g of quartz sand to a consistency of flour. The RNA was extracted with guanidinium thiocyanate and centrifuged through CsCl essentially as described in Sambrook et al., 1989, op. cit.. Poly A RNA was isolated from total RNA by chromatography on oligo dT cellulose.

15

**cDNA synthesis:** cDNA synthesis was carried out by means of a cDNA synthesis kit from Invitrogen according to the manufacturer's specifications. The DNA was adapted to the expression vectors by addition of a BstXI linker (Invitrogen) and size fractionated on an agarose gel. Only DNA larger than 5-600 bp was used in the library construction. The adapted cDNA was ligated into an appropriate vector cut with BstXI. Following test ligations (in order to determine the size of the library) the library was plated onto 50 agar plates. To each plate containing from approximately 500 to 5000 individual clones (dependent on the library size) was added 3 ml medium. The bacteria were scraped off, 1 ml glycerol was added, and stored at -80°C as 50 pools. The remaining 2 ml were used for DNA isolation. If the amount of DNA was insufficient to give the required number of yeast transformants (see below), large scale DNA was prepared from 500ml medium (TB) inoculated with 50 µl -80°C bacterial stock propagated over night.

**Construction of Yeast Libraries:** DNA from one or more pools was transformed into yeast as described below. To ensure that all the bacterial clones were tested in yeast a number of yeast

transformants 5 x larger than the number of bacteria clones in the original pools was set as a limit.

**Transformation of yeast:** The yeast strain used was yNG231. (MAT  
5 alpha, leu2, ura3-52, his4-539, pep4-delta 1, cir+). One colony  
was grown at 30°C overnight in 10 ml YPD (this culture can be  
stored for several days at 5°C).

10 10, 30, and 60 µl of this culture were added to 3 shaker flasks  
containing 100 ml YPD, and incubated with shaking overnight at  
30°C. The culture with an OD<sub>600</sub> closest to 0.3-0.4 was selected.  
The cells were harvested in 50 ml tubes in a Beckman centrifuge  
(speed 6, 10 minutes), the cells were resuspended in 2 x 5 ml  
H<sub>2</sub>O, centrifuged as described above, resuspended in 5 ml buffer  
15 containing 0.1 M LiAc, 10 mM Tris-Cl, 1 mM EDTA, pH 7.5, and  
centrifuged again. The cells were resuspended in 500 µl of the  
above buffer and incubated for 60 minutes at 30°C. 250 µg  
carrier DNA (sterile salmon-sperm DNA 10 mg/ml) was added and  
aliquots of 100 µl were prepared. The DNA to be transformed  
20 (approx. 5 µg) was added to the 100 µl aliquot, mixed gently,  
and incubated for 30 minutes at 30°C. 700 µl 40% PEG 4000, 0.1  
M LiAc, 10 mM Tris-Cl, 1 mM EDTA, pH 7.5 was added, and  
incubation was continued for 60 minutes at 30°C. The  
transformation mixture was subjected to heat shock for 5  
25 minutes at 42°C, spun briefly in a micro centrifuge,  
resuspended in 100-200 µl H<sub>2</sub>O, and plated on SC plates without  
uracil, followed by incubation for three days at 30°C.

**Preparation of carrier DNA:** 100 mg salmon-sperm DNA was weighed  
30 out and dissolved overnight in 10 ml 10 mM Tris-Cl, 1 mM EDTA,  
pH 7.5 (TE). The solution was then sonicated in a plastic  
container in ice water until it was no longer viscous. The  
solution was then phenole extracted and EtOH precipitated, and  
the pellet was washed and resuspended in 5 ml TE. The  
35 suspension was EtOH precipitated, and the pellet was washed and  
resuspend in 5 ml TE. The OD<sub>260</sub> was measured, and the suspension  
was diluted with TE to 10 mg/ml.

**Media:**

YPD: 10 g yeast extract, 20 g peptone, H<sub>2</sub>O to 810 ml.  
Autoclaved, 90 ml 20% glucose (sterile filtered) added.

- 5 10 x Basal salt: 66.8 g yeast nitrogen base, 100 g succinic acid, 60 g NaOH, H<sub>2</sub>O ad 1000 ml, sterile filtered.

SC-URA: 90 ml 10 x Basal salt, 22.5 ml 20 % casamino acids,  
9 ml 1% tryptophane, H<sub>2</sub>O ad 806 ml, autoclaved, 3.6 ml 5%  
10 threonine and 90 ml 20% glucose added.

SC-H agar: 7.5 g/l yeast nitrogen base without amino acids,  
11.3 g/l succinic acid, 6.8 g/l NaOH, 5.6 g/l casamino acids  
without vitamins, 0.1 g/l tryptophan and 20 g/l agar (Bacto).  
15 Autoclaved for 20 min. at 121°C. After autoclaving, 55 ml of  
a 22% galactose solution and 1.8 ml of a 5% threonine solution  
were added per 450 ml agar.

SC-H broth: 7.5 g/l yeast nitrogen base without amino acids,  
20 11.3 g/l succinic acid, 6.8 g/l NaOH, 5.6 g/l casamino acids  
without vitamins, 0.1 g/l tryptophan. Autoclaved for 20 min.  
at 121°C. After autoclaving, 10 ml of a 30% galactose solution,  
5 ml of a 30% glucose solution and 0.4 ml of a 5% threonine  
solution were added per 100 ml medium.

25

YNB-1 agar: 3.3 g/l KH<sub>2</sub>PO<sub>4</sub>, 16.7 g/l agar, pH adjusted to 7.  
Autoclaved for 20 min. at 121°C. After autoclaving, 25 ml of  
a 13.6% yeast nitrogen base without amino acids, 25 ml of a 40%  
glucose solution, 1.5 ml of a 1% L-leucine solution and 1.5 ml  
30 of a 1% histidine solution were added per 450 ml agar.

YNB-1 broth: Composition as YNB-1 agar, but without the agar.

CMC overlay gel: 1% agarose, 1% carboxymethyl cellulose in  
35 Tris-malate buffer, pH 7. The gel was boiled and then cooled  
to 55°C before the overlay was poured onto agar plates.

Oat spelt xylan overlayer gel: 1% agarose, 1% oat spelt xylan (Sigma Chemical Company) in Tris-malate buffer, pH 7. The gel was boiled and then cooled to 55°C before the overlayer is poured onto agar plates.

5

Construction of an Aspergillus expression vector: The vector pHD414 is a derivative of the plasmid p775 (described in EP 238 023). In contrast to this plasmid, pHD 414 has a string of unique restriction sites between the promoter and the terminator. The plasmid was constructed by removal of an approximately 200 bp long fragment (containing undesirable RE sites) at the 3'end of the terminator, and subsequent removal of an approximately 250 bp long fragment at the 5'end of the promoter, also containing undesirable sites. The 200 bp region was removed by cleavage with NarI (positioned in the pUC vector) and XbaI (just 3' to the terminator), subsequent filling in the generated ends with Klenow DNA polymerase +dNTP, purification of the vector fragment on gel and religation of the vector fragment. This plasmid was called pHD413. pHD413 was cut with StuI (positioned in the 5'end of the promoter) and PvuII (in the pUC vector), fractionated on gel and religated. The plasmid pHD 414 is shown in Fig. 2.

#### Example 1

25

A library from H. insolens consisting of approx. 300,000 individual clones in 50 pools was constructed.

DNA was isolated from 20 individual clones from the library and subjected to analysis for cDNA insertion. The insertion frequency was found to be >90 % and the average insert size was approximately 1400bp.

DNA was isolated from 10 pools from the Humicola library (2ml from the original plate). An aliquot was digested with restriction enzymes in order to excise the cDNA insert and analyzed by Southern blot using a 43kD cellulase cDNA probe

(the 43 kD enzyme is disclosed in PCT/DK91/00123) and a CBH 2 cDNA probe (the enzyme is disclosed in PCT/DK91/00124). Several bands were found to hybridize with the 43kD cellulase probe after a low stringency wash ( 2x SSC 65°C) in the 10 pools from the Humicola library. At higher stringency (0.1 x SSC, 75°C ) one band corresponding to the expected size for 43kD cellulase was detected in 5 out of 10 pools. Similar results were obtained with the CBH 2 probe. Here 10 out of 10 pools were found to have a band corresponding to the expected size for CBH 2. In addition, 4 pools contained bands with a higher molecular weight. These bands were seen even under stringent conditions, demonstrating that the library is of an appropriately high quality.

DNA from the Humicola library, pools 1-10, was transformed into yeast, and plates containing 20-25,000 colonies were obtained from each pool. The colonies were scraped off and stored in glycerol at -80°C.

Yeast cells from the library were spread onto YNB agar to a total of about 400,000 colonies. The number of colonies per plate varied from 50 to 500. After 4 or 5 days of growth, the agar plates were replica plated onto two sets of SC-H agar plates. These plates were then incubated for 2-4 days at 30°C before the two sets of agar plates were overlaid with a CMC indicator gel for detection of cellulase activity and oat spelt xylan indicator gel for the detection of xylanase and cellulase. After incubation overnight at 40°C, enzyme reactions were visualised with Congo Red. 10-15 ml of a 0.1% solution of Congo Red was poured onto the overlayer and removed after 10-20 min. The plates were then washed once or twice by pouring 10-15 ml of 2M NaCl onto the plates. The NaCl solution was removed after 15-25 min. Cellulase-positive colonies were identified on the plates with the CMC overlayer as colonies with colourless or pale red clearing zones on a red background. Xylanase-positive colonies identified on the plates with oat spelt xylan overlays as colourless or pale red clearing zones

on a red background. Cellulase-positive colonies were also identified on plates with oat spelt xylan overlayers as pale red or blue clearing zones on a red background.

- 5 Cells from enzyme-positive colonies were spread for single colony isolation on agar, and an enzyme-producing single colony was selected for each of the cellulase- or xylanase-producing colonies identified.
- 10 Each of the 133 cellulase-producing colonies and 147 of the xylanase-producing colonies were isolated. Some of these colonies were inoculated into 20 ml YNB-1 broth in a 50 ml glass test tube. The tube was shaken for 2 days at 30°C. The cells were harvested by centrifugation for 10 min. at 3000 rpm.
- 15 The cells were resuspended in 1 ml 0.9 M sorbitol, 0.1 M EDTA, pH 7.5. The pellet was transferred to an Eppendorf tube, and spun for 30 seconds at full speed. The cells were resuspended in 0.4 ml 0.9 M sorbitol, 0.1 M EDTA, 14 mM  $\beta$ -mercaptoethanol.
- 20 100  $\mu$ l 2 mg/ml Zymolase was added, and the suspension was incubated at 37°C for 30 minutes and spun for 30 seconds. The pellet (spheroplasts) was resuspended in 0.4 ml TE. 90  $\mu$ l of (1.5 ml 0.5 M EDTA pH 8.0, 0.6 ml 2 M Tris-Cl pH 8.0, 0.6 ml 10% SDS) was added, and the suspension was incubated at 65°C
- 25 for 30 minutes. 80  $\mu$ l 5 M KOAc was added, and the suspension was incubated on ice for at least 60 minutes and spun for 15 minutes at full speed. The supernatant was transferred to a fresh tube which was filled with EtOH (room temp.) followed by thorough but gentle mixing and spinning for 30 seconds. The
- 30 pellet was washed with cold 70% EtOH, spun for 30 seconds and dried at room temperature. The pellet was resuspended in 50  $\mu$ l TE and spun for 15 minutes. The supernatant was transferred to a fresh tube. 2.5  $\mu$ l 10 mg/ml RNase was added, followed by incubation at 37°C for 30 minutes and addition of 500  $\mu$ l
- 35 isopropanol with gentle mixing. The mixture was spun for 30 seconds, and the supernatant was removed. The pellet was rinsed with cold 96% EtOH and dried at room temperature. The DNA was

dissolved in 50  $\mu$ l water to a final concentration of approximately 100  $\mu$ l/ml.

The DNA was transformed into E.coli by standard procedures.  
 5 Two E. coli colonies were isolated from each of the transformations and analysed with the restriction enzymes HindIII and XbaI which excised the DNA insert. DNA from one of these clones was retransformed into S. cerevisiae strain JG169 (MAT $\alpha$ ; ura 3-52; leu 2-3, 112; his 3-D200; pep 4-113;  
 10 prcl::HIS3; prb1::LEU2) and rescreened for enzyme activity.

The DNA sequences of several of the positive clones were partially determined. The partial DNA sequences are shown in Sequence Listings SEQ ID#7-15. Based on the DNA sequence, the  
 15 clones were classified as follows:

#### Endocellulases:

|             |  |   |
|-------------|--|---|
| CMC 1:      | C3, 26, 27, XY33, XY46                               | 250 amino acids (SEQ ID#7)                        |
| 20 CMC 4:   | C46, 47, 50, 51, 54,<br>101, 102, 103, 104           | ~1400 bp (the enzyme of the invention) (SEQ ID#8) |
| CMC 5:      | XY49   | ~1050 bp (SEQ ID#9)                               |
| CMC 6:      | C49  | ~1000 bp (SEQ ID#10)                              |
| 25 CMC 38K: | C13  | (SEQ ID#11)                                       |
| CMC EG1:    | C6, 11, 15, 16, 17, 21,<br>22, 23, 25, XY34, 41, 145 | (SEQ ID#12)                                       |

#### Xylanases:

|    |        |   |                   |
|----|--------|---|-------------------|
| 30 | XYL 1: | XY30, 31, 40, 42, 101, 102,<br>110, 117, 119, 123, 125,<br>136, XY56, 60, 137               | 22 kD (SEQ ID#13) |
| 35 | XYL 2: | XY103, 104, 107, 108, 109,<br>113, 114, 118, 120, 121, 124,<br>126, 128, 130, 134, 142, 143 | (SEQ ID#14)       |
|    | XYL 3: | XY115, 116, 132, 146  | (SEQ ID#15)       |

In order to express the genes in Aspergillus the cDNA insert is isolated from one or more representatives of each family and cloned into the vector pHD414 which is transformed into A. oryzae or A. niger according to the general procedure described below.

Transformation of Aspergillus oryzae or Aspergillus niger (general procedure)

10

100 ml of YPD (Sherman et al., Methods in Yeast Genetics, Cold Spring Harbor Laboratory, 1981) is inoculated with spores of A. oryzae or A. niger and incubated with shaking at 37°C for about 2 days. The mycelium is harvested by filtration through miracloth and washed with 200 ml of 0.6 M MgSO<sub>4</sub>. The mycelium is suspended in 15 ml of 1.2 M MgSO<sub>4</sub>. 10 mM NaH<sub>2</sub>PO<sub>4</sub>, pH = 5.8. The suspension is cooled on ice and 1 ml of buffer containing 120 mg of Novozym<sup>®</sup> 234, batch 1687 is added. After 5 minutes 1 ml of 12 mg/ml BSA (Sigma type H25) is added and incubation with gentle agitation continued for 1.5-2.5 hours at 37°C until a large number of protoplasts is visible in a sample inspected under the microscope.

The suspension is filtered through miracloth, the filtrate transferred to a sterile tube and overlaid with 5 ml of 0.6 M sorbitol, 100 mM Tris-HCl, pH = 7.0. Centrifugation is performed for 15 minutes at 100 g and the protoplasts are collected from the top of the MgSO<sub>4</sub> cushion. 2 volumes of STC (1.2 M sorbitol, 10 mM Tris-HCl, pH = 7.5. 10 mM CaCl<sub>2</sub>) are added to the protoplast suspension and the mixture is centrifugated for 5 minutes at 1000 g. The protoplast pellet is resuspended in 3 ml of STC and repelleted. This is repeated. Finally the protoplasts are resuspended in 0.2-1 ml of STC.

100 µl of protoplast suspension is mixed with 5-25 µg of the appropriate DNA in 10 µl of STC. Protoplasts are mixed with p3SR2 (an A. nidulans amdS gene carrying plasmid). The mixture



is left at room temperature for 25 minutes. 0.2 ml of 60% PEG 4000 (BDH 29576). 10 mM  $\text{CaCl}_2$  and 10 mM Tris-HCl, pH = 7.5 is added and carefully mixed (twice) and finally 0.85 ml of the same solution is added and carefully mixed. The mixture is left  
5 at room temperature for 25 minutes, spun at 2500 g for 15 minutes and the pellet is resuspended in 2 ml of 1.2 M sorbitol. After one more sedimentation the protoplasts are spread on the appropriate plates. Protoplasts are spread on minimal plates (Cove Biochem. Biophys. Acta 113 (1966) 51-56)  
10 containing 1.0 M sucrose, pH = 7.0, 10 mM acetamide as nitrogen source and 20 mM CsCl to inhibit background growth. After incubation for 4-7 days at 37°C spores are picked and spread for single colonies. This procedure is repeated and spores of a single colony after the second reisolation is stored as a  
15 defined transformant.

#### Example 2

Cellulase type 4 clones C46 and C51 and a 43 kD cellulase  
20 control clone (obtained by transforming yeast strain JG169 with PYHD17 carrying a DNA sequence coding for the 43 kD cellulase [isolated as described in PCT/DK91/000123]) were inoculated in 100 ml test tubes with 15 ml YNB-1 broth. The tubes were agitated at 30°C for 2 days. 5 ml of broth from each tube were  
25 then used as seed material for shake flasks containing 100 ml SC-H broth. The shake flasks were agitated for 4 days at 30°C. The cells from 20 ml of broth were collected by centrifugation and mixed with 1-2 ml 0.1 M sodium phosphate buffer, pH 7, and 3.3 g of glass beads (420-500  $\mu\text{m}$  in diameter) in 10 ml glass  
30 test tubes. The crude cell extracts were collected after about 8 minutes of agitation by means of a IKA vibrax VXR (available from IKA Labortechnik).

The cellulase activity of the crude cell extracts from the  
35 yeast clones C46, C51 and 43 kD were measured under different conditions by the size of the clearing zones formed in CMC containing gels.

CMC gel: CMC overlayer gel as described above.

CMC LAS gel: 2% agarose, 1% CMC in 50 mM sodium phosphate buffer, pH 7, boiled and mixed with one volume of 0.12% LAS.

5

The cellulase activity was measured by adding 15  $\mu$ l crude cell extract to 4 mm (diameter) holes in the gel. The crude cell extracts were diluted with one volume of 0.12% LAS before addition to the CMC LAS gel and with one volume of water before addition to the CMC gel. The clearing zones were then visualised after 18 hours of incubation at 40°C by staining with Congo Red as described above.

The results are shown in the following table.

15

| Clone   | C46 | C51 | 43 kD |
|---------|-----|-----|-------|
| CMC     | 14  | 14  | 17    |
| CMC LAS | 15  | 13  | 0     |

20

Activities are shown as mm clearing zones.

25

It appears from the table that the enzyme produced by C46/C51 is LAS resistant.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANT:  
(A) NAME: Novo Nordisk A/S  
(B) STREET: Novo Alle  
(C) CITY: Bagsvaerd  
10 (E) COUNTRY: Denmark  
(F) POSTAL CODE (ZIP): DK-2880  
(G) TELEPHONE: +45 4444 8888  
(H) TELEFAX: +45 4449 3256  
15 (I) TELEX: 37304
- (ii) TITLE OF INVENTION: A Method of Cloning Proteins in Yeast
- (iii) NUMBER OF SEQUENCES: 15
- 20 (iv) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (EPO)
- (2) INFORMATION FOR SEQ ID NO: 1:
- 30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 93 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:  
40 (A) ORGANISM: Humicola insolens

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

45 TGGCAGCAGT GGGTGGGCT TGGCTTCTCG GGCCTCTAGT CCTGCTGTC CGGTACACG 60  
TGGTGTACT TGAAGACTG GTACAGCCAA TGC 93

## (2) INFORMATION FOR SEQ ID NO: 2:

- 50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 92 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
55 (D) TOPOLOGY: linear

26

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Humicola insolens

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CAGCGCAGCC GACGAGTTA OGGACAACAC AACAAAGACC AGGGCAACAT OGACAACAAG 60  
10 GTGAGCCCCG GCTGCCACTT CAACCACTCC GG 92

(2) INFORMATION FOR SEQ ID NO: 3:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 132 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Humicola insolens

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CCAAGGCGAA GTTCAAGTGG TTGGCATCAA OCAGTCTGC GCTGAGTTGG GCAAGGAGAG 60  
30 TATCGGGCTA TGGGCAAGCA CTTACTTCTT TOGOGAAGTC GTGATTCAA GGGCACATCA 120  
ATGTTGGCTT CA 132

35 (2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 69 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: cDNA

45 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Humicola insolens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

50 CTGAGTGAA OGTAACCAAC AACCACTTGG CCGTAGGAC OGAGAACAAG CTGTGTACCA 60  
GATGCATCA 69

55 (2) INFORMATION FOR SEQ ID NO: 5:

27

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 120 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
5 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:  
10 (A) ORGANISM: *Humicola insolens*
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
- 15 GGAAGGTCOG GCAAGAGCAC GGCTGGGTC AGCACCAGG TOGGGCTTCA GOGGTCATT 60  
GGGOGACCA ACTGGCTCAG GCAAAAGGC AAGGTGGAC TGCTGGGAC TTGOGGGC 120
- 20 (2) INFORMATION FOR SEQ ID NO: 6:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 120 base pairs  
(B) TYPE: nucleic acid  
25 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:  
30 (A) ORGANISM: *Humicola insolens*
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
- 35 GCCAAGTGGG TTGOCAGCA GGCCATGAG GGCATGCTGA ACCACCTCCA GGAGAATAGC 60  
GATGTCGGA CAGGTGGGCT CTGGTGGGCG GGAGGCGGCT GGTGGGGTTG ACTATATCTA 120
- 40 (2) INFORMATION FOR SEQ ID NO: 7:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1027 base pairs  
45 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- 50 (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Humicola insolens*
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CATOGCCTTA TAACCACAGC TCTACTGCAG ACCTTGTOCA ATTTCTOGGA TCACGGCCAT 60  
 GCICAAGTCT GCTCTOCTOC TOGGGCOOCC GGOOGTTTOG GTTCAGTOGG CTTOGATOCC 120  
 5 GAACATCOOG GCAAACCTTG AGCCTOGCCA GATTGCGTOG CTCTGTGAGC TGTAOGGCTA 180  
 CTGGTOGGGC AATGGCTATG AGCTGTTGAA CAACCTCTGG GGCAAGGACA COGOCACATC 240  
 OGGCTGGCAA TGCACCTACC TOGAOCCGAC CAACAACGGT GGCATTTCAGT GGAGCAOCC 300  
 10 GTGGGAGTGG CAGGGOGCTC OGGACAAOGT CAAGAGCTAC CCTATGTTG GCAAGCAGAT 360  
 CCAGOGGGC OCAAGATCA GOGACATCAA CAGCATGOGC AOGTOGGTGT OGTGGAOGTA 420  
 15 OGATOGGACC GACATCOGTG CCAATGTGCG TTATGATGTC TTCAOGGCTC GTGATCOGGA 480  
 CCATOOCAAC TGGGGOGGCG ACTACGAGCT CATGATCTGG CTGGOOOGCT ATGGOGGCGAT 540  
 CTACCCCATC GGCAOGTTCC ACAGCCAGGT CAACCTTGCT GGTGCTACCT GGGATCTCTG 600  
 20 GACTGGCTAC AAOGGCAACA TGGTGTCTCA CAGCTTCCTC CCCCCGTCCG GOGACATTGG 660  
 TGACTTCAGC TGGACATCA AGGACTTCTT CAACTACCTT GAGOGCAACC ATGGCTACCC 720  
 25 GGCAOGGGAG CAGAATCTGA TGTCTAACA AGTTGGAACC GAGTGCCTCA OGGGOGGTCC 780  
 GGCCAGGTTT AOGTGCAGGG ACTTCAGGGC TGACCTGTTG TAAGAGGGGT CATTTGAGTA 840  
 GGGTGTACTT GOCAGGAAGC GGAGAGGAAC AAGGTAGATA TAATGACAGA CAAGTCATTG 900  
 30 GATGCAGTAA ACAGGCTTGG TCCITCATGG GCAAGACTTG CAAGTGOCTA GACAACAGTG 960  
 CTCAGGGGAC ACAGGGGAOGA AGATGCATCC TCCAAACTC AGAAATOGGT AGCAATOGCC 1020  
 35 TTATAGAG 1027

## (2) INFORMATION FOR SEQ ID NO: 8:

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 872 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Humicola insolens*

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AAAGCCTGAA CACTATTACC ATGTTGCACA GTGTCTCTTC GOGACTGGAG 60  
 55 OGCTOGCCA GGGGTGTCAT GGCAGCAGTG TGGTGGGTTT GGCTTCTOGG GCTCTACGTC 120

CTGTGTGTCC GGTTACAGT GGTGTACTT GAAAGACTGG TACAGCCAAT GCCAGGCGAG 180  
 CCGAGAGAGT TAAGGACAAC ACAACAAGC CAGGGCAACA TOGACAACAA GTTACAGCCCC 240  
 5 GGCTGCCACT TCAACCACTC CGGCCAAGGC GAAGTTCAAG TGGTTGGCAT CAACCACTCC 300  
 TGGCTGAGT TOGGCAAGGA GAGTATCOGG CTATGGGCAA GCACTTACTT CCTTCGGGAC 360  
 GTGGTGGATT CAAGGCGACA TCAATGGTGG CTTACATGT CCGNNNNGGC AACTCTGGAC 420  
 10 GGCTGAGGTG AAGTGAACA ACAACAATT GGCCTAGOG ACGAGAACA AGCTGTGTAC 480  
 CAGATGCATC AGTACCTOGA CTGGGAGGT CCGGCAOGAG CAGGGGCTGC GTCAGCACCC 540  
 15 AGGTGGGCT TCAGGGGTC ATTGGGGGA CCACTGGCT CAGGCAAAAC GGCAAGGTTG 600  
 GACTGCTGCG GACTTGGCG GGGGCAAGT GGGTTTGOCA GCAGGCAAT GAGGGCATGC 660  
 TGAACCACT CCAGGAGAAT AGGATGTCT GGACAGGTGC GCTCTGGTGG GGGGAGGCG 720  
 20 GGTGGTGGG TGAATATATC TAGTGGTTG AACCTCTTG GGTATGGCT ACACCTACTA 780  
 CAATTCCTT CTCAAGAAAT AGTGGCATA GGTGCTATAA GACGTGGTC CTGACTCAAG 840  
 25 AGGGTTTGAC AGGAAGCAG CCTGAGGCT TT 872

## (2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:  
 30 (A) LENGTH: 368 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 35 (ii) MOLECULE TYPE: cDNA  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Humicola insolens*  
 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GTTGAAGGCG CTGAACAGAG GAAGTCAAGT CCGAAAAATG TCCAGGGCTA CGGAGTATGT 60  
 45 ACAGAATATC CACAACAAA CAAAGTCAA TTCACCTGCG ATCCCAACAT CACAACCCCT 120  
 TCAGGCTGGG TOGGAAACAA CTACCTAGCC AGCCTCTCTC CATTTTCCAT CTCCTAACAC 180  
 CAACCCCTT CCAATCTCTG CCGCTTAAAT CGGGTTGACC CGAATGATC CAGATCGCA 240  
 50 AACCAACCT GCAACCAATC CGCTGGGCG GCGCTGGGTG AAGAAAGACT TGGCGGTCT 300  
 CGTTGCGGGA GCAGGCACTC GAACAGTAGC TGGCGCGGT GGAAGTTGC CGTTGTAGTA 360  
 55 CCATAGGT 368

## (2) INFORMATION FOR SEQ ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 720 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

10

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Humicola insolens*

## 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

NAAAGGCAOC AAGGTGAOCC OGTCACCTOG GGOGAGTGGG AGAOGATCOG CATCAOOGAG 60  
AACCCTGGT OGCAOCCGCC OGTCACCGAC GTGACCTOGC OGGOCATGAC GTGCTACGAG 120  
20 AAGACGCCOG GCCAGOGGGC CCNAAGACGG TCAAGTGGG GGCCGGGGOG ACOGTCAOCT 180  
TCAOCTCTA CAOCCAGTGG GGCAOCCGG GGGCTGCACT TCTAOCCTGGC CAAGGTGCOG 240  
25 OGGCAAGACG GGCCGACGT TTGAOCCGAA GGGGGCCGTG TGGTCAAGA TTTCACAGGA 300  
CGGCTGGTG GTTGGGACAG CTGGTTCACC TGGCATGCT TGGCAAGAAG AGTCTGTCA 360  
ATNNNNNNAC CTGGCTACG TTGGCAAGAA GAAGTCTGT CAATCCCCOG TGCTCCAGG 420  
30 ACGGCGAGTA CCGTGGGGG TOGAGCACAT TGCTGCACA GGCGGAGOG TGGGGGTGC 480  
GCAGCTCTAC ATTGTGTGG OGCAATCAA OGTCACGGG GCAOCCGCAC GCTCAACCCG 540  
35 GGCCAGCTOG TCTGTTCOC GGGGGCTAC AAGCCCAOC ACCOGGGCAT CCGTTCACG 600  
CTCTACTGGC OGGGGGGAC GCAGTACATC AACCCGGTC OGGGGGGGT GAAGTCTGA 660  
GTTCAGTTC ATGAGTACTC CAATGAAGGT TGCGGGGGG OGAGGGTAGG TOGATAGTTT 720  
40

## (2) INFORMATION FOR SEQ ID NO: 11:

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 724 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

50

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Humicola insolens*

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:



OCTAGGTGCG CCAOCATGCG OGTTTCTCTT GCTCTCCTCG OCTACCTGCT CAGGCGCGCC 60  
 CCGGCTCGC CGTCCCGGA GCTGAGGCC CGGCAGTCGG GCAACCCCTT CTCGGGCGCC 120  
 5 ACCCTGCTGG TCAACTGGA CTATAGCAGC AAGCTGACC AGACGGGCA GGCTTTTCTT 180  
 GTCCGCGGC GACCAGACA AGCTGCCAA GGTCAAGTAC GTCCAGGAGA AGGTTGGCAC 240  
 CTTTCTATTG GACTTCCAAC ATCTTCTCTC TGCGCAGCAC TGAAGTTGCG ATCCAGAATG 300  
 10 CGGCGGCGCA AGGCGGCGG AGAACCCCAT GTGGGTCTC GTCTGTACA ACCTCCCGA 360  
 CCGGACTGC AGGAGCGGG CAGTACCTCT GGGAAGTTA AGCTCTCCA GAAAGGCGTG 420  
 15 AACGGGTACA AGAAGAGTA CGTCAACCG TTGCGCCAGA AGCTCAAGGC CGGTCGAC 480  
 GTGCAGTTGG CGTCATCTT CGAGCCGAT GGCATGGCA ACATGGTCAC GGGCACCAGC 540  
 GCCTTCTGCG GCAAGCGCG CGGCCCTCAG AGGAGGCCAT CGGCTATGCT ATCTCTCTC 600  
 20 GGCTGGGCG ATAAGCTGA GCAACTGCG CAGGAGGTGC CACCATCTC CAAAAGGCG 660  
 GTAACACGC AAGATGGGG CTTCTCAGCA AGTTCCAAC TACAACCTAT TCAAGACAAC 720  
 25 CGCG 724

## (2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:  
 30 (A) LENGTH: 71 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 35 (ii) MOLECULE TYPE: cDNA  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Humicola insolens*

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

OCTCCTTCCA GTCCTTGAGT TCCTTGGAC CTGCAGGTC CTGAACAACT CGCTCTAGCT 60  
 45 CAACAACCAT G 71

## (2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:  
 50 (A) LENGTH: 572 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 55 (ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Humicola insolens*

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

|    |   |     |
|----|---|-----|
|    | GGCGGTGGCC GCTTCITCAG TTGTGTAAGA TCATOCAGCA ACTGGCAGTT CAOCATGGTC | 60  |
|    | TGCTCAAGT CTGTCTCTGC GGCAGCCAG GCTGTGAGCT CTGCAATTGC TGCCCCITTT   | 120 |
| 10 | GACTTGGTTC CTGGGACAA CTGAAGGCC CTTCAGGCTC GGCAGGTGAC CCCCAGGCC    | 180 |
|    | GAGGGCTGGC ACAAGGGCTA CTTCTACTOG TGGTGGTGG AOGGOGGAGG CCAGGTTCAG  | 240 |
| 15 | TACACCAACC TGAGGGCAG CCGCTACCAG GTCAGATGGC GTAACACCG CAACTTGGTC   | 300 |
|    | GGTGGTAAGG GTTGAACCC GGAACCGGC CGCAOGATCA ACTAOGGOGG CTACTTCAAC   | 360 |
|    | CCCCAGGGCA AOGGCTACCT GGCCTCTAC GGCTGGACCC GCAACCGCT CGTOGAGTAC   | 420 |
| 20 | TATGTCATOG AGTGTACCG CAOGTACAAT CCGGCAGCC AGGCTCAGTA CAAGGGCACA   | 480 |
|    | TTCTATACOG ACGGCGATCA GTATGACATC TTTGTGAGCA CCGTCACAA CCAGCCAGC   | 540 |
| 25 | ATCAAGGCAC CCGAAGTCC AGCTAGTACT GG                                | 572 |

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:  
30 (A) LENGTH: 173 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Humicola insolens*

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

|    |   |     |
|----|---|-----|
|    | AAAGTAGATC GTCCTGCTCC CTAGAAACCA GTCACATATT CACAATGGGT TCTATOGCTC | 60  |
| 45 | TGCTCTTTC TGCGGCTCCG GGCTCTCTCG CCAGTCCCA GCTCTGGGC AGTGGGGTGG    | 120 |
|    | CATGGCTGG AAGGCTTAC GACTTGGTTC TGGGGGCTA CCTGCACCAA GAT           | 173 |

(2) INFORMATION FOR SEQ ID NO: 15:

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 214 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
55 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Humicola insolens*

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AAAGCOGCAG CGATCAAGAA CCAAGCAGT CTGTCAAAT GGTGGCTTC TGTCCCTCT 60  
10 TCTCGGTCG TCCATGGCC GCAAGGCTC GCGCGGCCC GGTGTAGCTG CCGGCGATG 120  
ACCTGAACAA GGTTCAGAC TACCCCGAG AGGCTACCG GCACTCACA CCGCTACATC 180  
15 TTCTCCTTCT GGACTGAAG TCTGGCGAA CGTC 214

## CLAIMS

- 5 1. A method of screening for a DNA sequence coding for a protein of interest, the method comprising
- (a) cloning, in suitable vectors, a DNA library from an organism suspected of producing one or more proteins of  
10 interest,
- (b) transforming suitable yeast host cells with said vectors,
- (c) culturing the host cells under suitable conditions to  
15 express any protein of interest encoded by a clone in the DNA library, and
- (d) screening for positive clones by determining any activity of a protein expressed in step (c).
- 20 2. A method according to claim 1, wherein the DNA library is a cDNA library prepared from the mRNA of an organism suspected of producing one or more proteins of interest.
- 25 3. A method according to claim 1 or 2, wherein positive clones isolated in step (d) are subjected to rescreening, reisolation and recloning.
4. A method according to claim 1 or 2, wherein the organism  
30 suspected of producing one or more proteins of interest is a eukaryotic organism.
5. A method according to claim 4, wherein the eukaryotic organism is a fungus.
- 35 6. A method according to claim 4, wherein the eukaryotic organism is a plant.

7. A method according to any of claims 1-6, wherein the protein of interest is an enzyme.

8. A method according to claim 1, wherein the yeast host cell is a strain of Saccharomyces cerevisiae, Schizosaccharomyces pombe, Hansenula, Pichia, Yarrowia lipolytica or Kluyveromyces lactis.

9. A process for producing a protein of interest in a heterologous host cell, the process comprising transforming a suitable heterologous host cell with a DNA sequence coding for a protein of interest, which DNA sequence has been isolated by the method of claim 1, culturing the transformed cells under suitable conditions to express the protein, and recovering the expressed protein from the culture.

10. A process according to claim 9, wherein the DNA sequence coding for the protein of interest has been isolated by the method of claim 3.

11. A process according to claim 9 or 10, wherein the protein of interest is an enzyme.

12. A process according to any of claims 9-11, wherein the host cell is a strain of Aspergillus, for instance a strain of Aspergillus oryzae or Aspergillus niger.

13. An enzyme which exhibits cellulase activity, and which has the following characteristics

(a) the DNA sequence encoding the enzyme has been isolated from a DNA library of Humicola insolens,

(b) said DNA sequence comprises at least one of the following partial sequences

36

(i) TGGCAGCAGT GTGGTGGCGT TGGCTTCTCG GGCTCTACGT  
 CCTGTGTGTC CGGTTACACG TGC GTGTACT TGAACGACTG  
 GTACAGCCAA TGC  
 (SEQ ID#1)

5 (ii) CAGCGCAGCC GACGACGTTA CGGACAACAC AACAAACGACC  
 AGGGCAACAT CGACAACAAG GTCAGCCCCG GCTGCCACTT  
 CAACCACTCCG G  
 (SEQ ID#2)

10 (iii) CCAAGGCGAA GTTCAAGTGG TTGGCATCAA CCAGTCCTGC  
 GCTGAGTTCG GCAAGGAGAG TATCCGGCTA TGGGCAAGCA  
 CTTACTTCCT TCGCGACGTC GTCGATTCAA GCGCACATCA  
 ATCGTGGCTT CA (SEQ ID#3)

15 (iv) CTGACGTGAA CGTGACCAAC AACAACTTGG CCGTAGCGAC  
 CGAGAACAAG CTGTGTACCA GATGCATCA (SEQ ID#4)

20 (v) GGACGGTCCG GCACGAGCAC GGCCTGCGTC AGCACCCAGG  
 TCGGCCTTCA GCGCGTCATT GGC GCGACCA ACTGGCTCAG  
 GCAAAACGGC AAGGTTGGAC TGCTCGCGAC TTGCCGCGGC (SEQ  
 ID#5)

25 (vi) GCCAAGTGGG TTTGCCAGCA GGCCATTGAG GGCATGCTGA  
 ACCACCTCCA GGAGAATAGC GATGTCTGGA CAGGTGCGCT  
 CTGGTGGGCG GGAGGCCCGT GGTGGGGTTG ACTATATCTA (SEQ  
 ID#6)

- (c) the enzyme comprises a cellulose-binding domain, and
- 30 (d) the enzyme exhibits endocellulase activity in the presence of linear alkyl benzene sulfonate.

14. An enzyme according to claim 13, a crude extract (15  $\mu$ l)  
 35 of which diluted with one volume of 0.15% linear alkyl benzene sulfonate and added to a 2% agarose gel containing 1% carboxymethyl cellulose in 50 mM sodium phosphate buffer,

pH 7, mixed with one volume of 0.15% linear alkyl sulfonate forms a clearing zone in said agarose gel after 18 hours of incubation, which clearing zone is equal to (less 3 mm) the clearing zone formed in a similar carboxymethyl cellulose gel not containing any linear alkyl benzene sulfonate, provided that the concentration of enzyme in the extract is such that a clearing zone of at least 10 mm is formed in a carboxymethyl cellulose gel (with no linear alkyl benzene sulfonate) under the conditions specified above.

10

15. A detergent additive comprising an enzyme exhibiting cellulase activity according to any of claims 13-14, preferably in the form of a non-dusting granulate, stabilised liquid or protected enzyme.

15

16. A detergent additive according to claim 15, which further comprises one or more other enzymes such as a protease, amylase, lipase or peroxidase.

20 17. A detergent composition comprising an enzyme exhibiting cellulase activity according to any of claims 13-14.

18. A detergent composition according to claim 17, which further comprises a linear alkyl benzene sulfonate surfactant.

25

19. A detergent composition according to claim 17 or 18, which further comprises one or more other enzymes such as a protease, amylase, lipase or peroxidase.

30

1/2

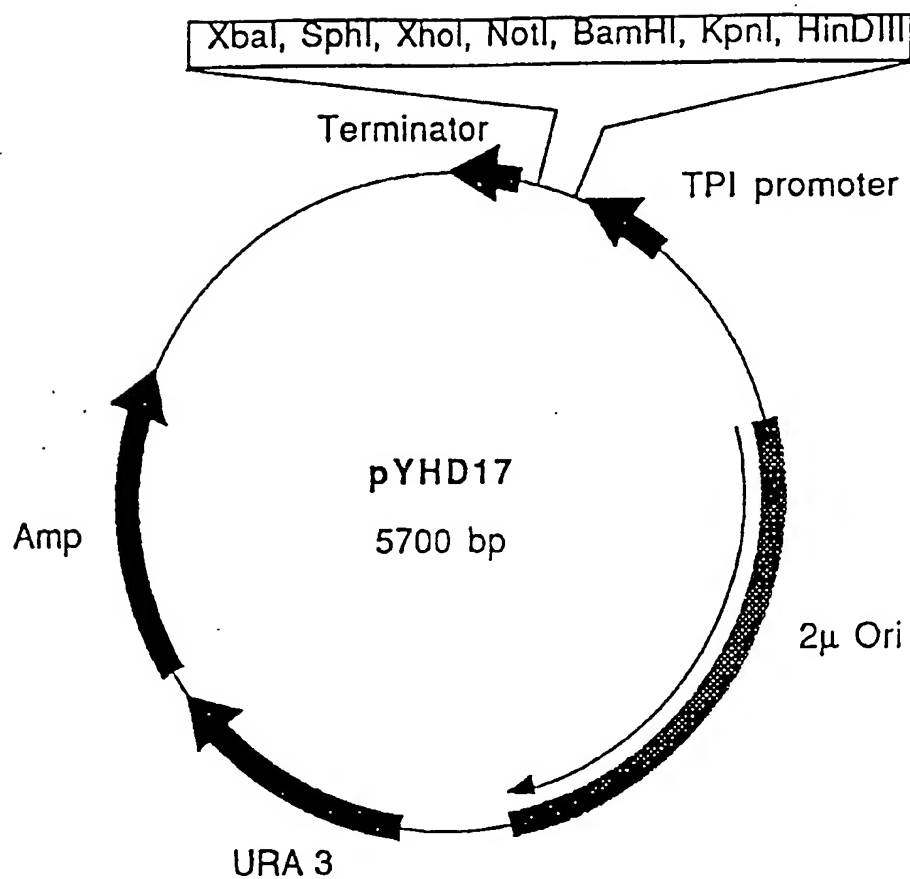


Fig. 1



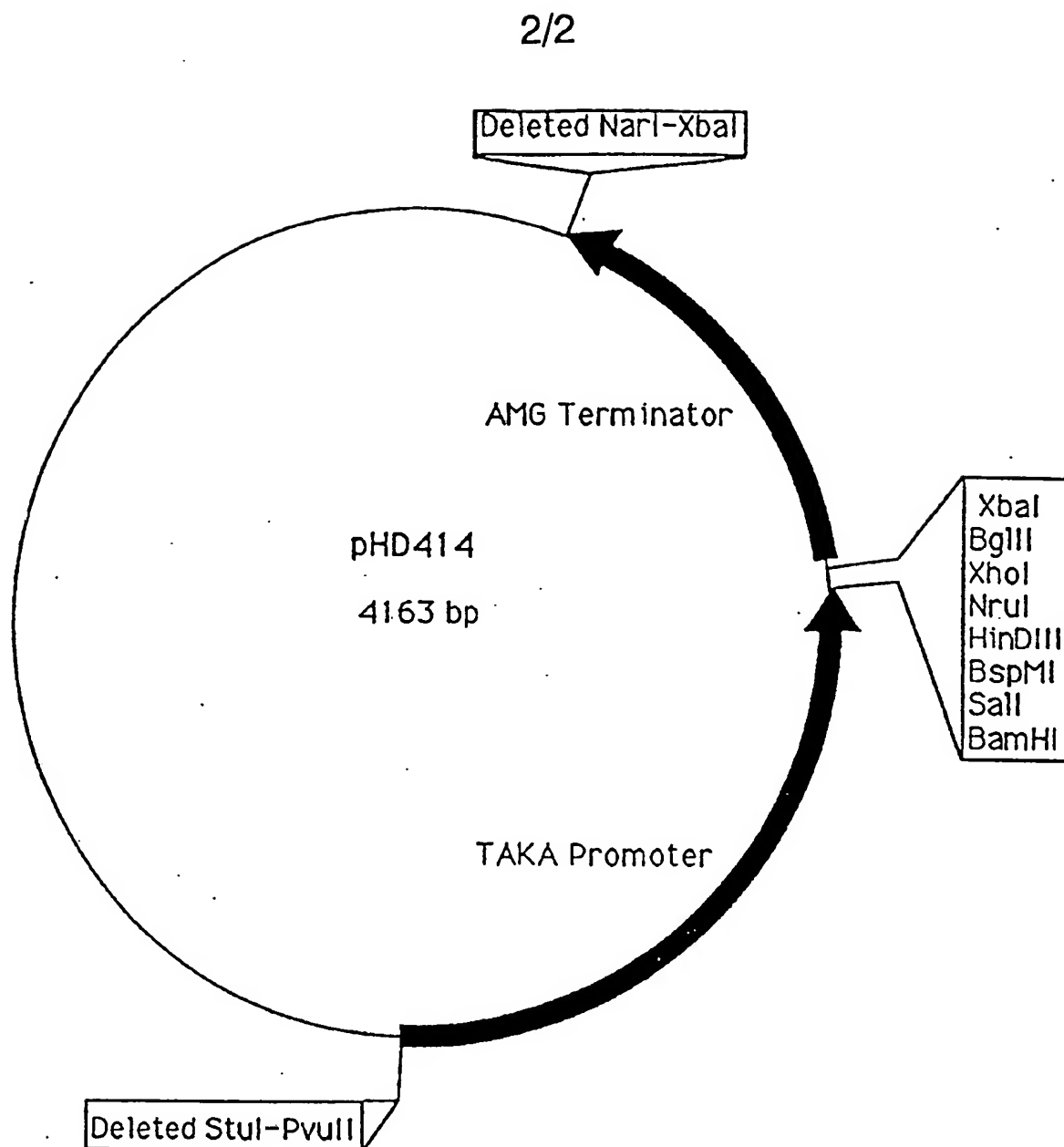


Fig. 2

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 92/00360

## A. CLASSIFICATION OF SUBJECT MATTER

IPC5: C12N 15/81, C12N 1/16, C12N 9/42

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC5: C12N, C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, CA, EMBL

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|-----------|---|-----------------------|
| P,X       | WO, A1, 9201069 (TRANSKARYOTIC THERAPIES, INC.),<br>23 January 1992 (23.01.92), page 9,<br>line 14 - line 30, claims 1-2<br>--  | 1-12                  |
| X         | PROC.NATL.SCI., Volume 80, July 1983,<br>Gary L. McKnight et al, "Selection of functional<br>cDNAs by complementation in yeast"<br>page 4412 - page 4416<br>--  | 1-12                  |
| X         | SCIENCE, Volume 236, 1987, David T. Burke et al,<br>"Cloning of Large Segments of Exogenous DNA into<br>Yeast by Means of Artificial Chromosome Vectors",<br>page 806 - page 812, page 808 right column line<br>10-12<br>-- | 1-12                  |

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

## \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

16 April 1993

20 Apr 1993

Name and mailing address of the ISA/  
Swedish Patent Office  
Box 5055, S-102 42 STOCKHOLM  
Facsimile No. +46 8 666.02 86

Authorized officer

Yvonne Siösteen  
Telephone No. +46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 92/00360

| C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |   |                       |
|---|---|-----------------------|
| Category*   | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
| X   | Chemical Abstracts, Volume 114, No 5,<br>4 February 1991 (04.02.91), (Columbus, Ohio, USA),<br>Okazaki, Koei et al, "High-frequency<br>transformation method and library transducing<br>vectors for cloning mammalian cDNAs by<br>transcomplementation of Schizosaccharomyces pombe",<br>page 202, THE ABSTRACT No 37151y, Nucleic Acids<br>Res. 1990, 18 (22), 6485-6489<br><br>-- | 1-12                  |
| A   | THE EMBO JOURNAL, Volume 8, No 12, 1989,<br>David P. Gearing et al, "Expression cloning of a<br>receptor for human granulocyte-macrophage<br>colony-stimulating factor" page 3667 - page 3676<br><br>--   | 1-12                  |
| A   | WO, A1, 8909259 (NOVO INDUSTRI A/S),<br>5 October 1989 (05.10.89), see claim 9 and page 3,<br>page 7 and example 1<br><br>--  | 13-19                 |
| A   | WO, A1, 9117243 (NOVO NORDISK A/S),<br>14 November 1991 (14.11.91)<br><br>--  | 13-19                 |
| A   | US, A, 4435307 (NOVO INDUSTRI A/S), 6 March 1984<br>(06.03.84)<br><br>--  | 13-19                 |
| A   | Chemical Abstracts, Volume 110, No 19, 8 May 1989<br>(08.05.89), (Columbus, Ohio, USA), Rao U.<br>Subrahmanyaswara et al., "Purification and<br>characterization of a beta-glucosidase and<br>endocellulase from Humicola insolens", page 358,<br>THE ABSTRACT No 168875c, Indian J. Biochem.<br>Biophys. 1988, 25 (6), 687-694, (e)<br><br>--                                      | 13-19                 |

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 92/00360

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|-----------|---|-----------------------|
| A         | Chemical Abstracts, Volume 71, No 21,<br>24 November 1969 (24.11.69), (Columbus, Ohio, USA),<br>Ramabadran R. et al., "Cellulase of Humicola<br>insolens", page 33, THE ABSTRACT No 98499h, Indian<br>J. Exp. Biol. 1969, 7 (3), 186-187, (e)<br><br>--                                   | 13-19                 |
| A         | Chemical Abstracts, Volume 109, No 23,<br>5 December 1988 (05.12.88), (Columbus, Ohio, USA),<br>Hayashida Shinsaku et al., "Cellulases of Humicola<br>insolens and Humicola grisea", page 295,<br>THE ABSTRACT No 207112c, Methods Enzymol. 1988,<br>160, 323-332, (e)<br><br>--<br>----- | 13-19                 |

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 92/00360

**Box I** Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II** Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

The inventions claimed are composed of the following two different inventions:

- I. Claims 1-12 directed to a method of screening for a DNA sequence coding for a protein of interest and a process for producing the protein.
  - II. Claims 13-19 directed to an enzyme and its use. The "special technical features" of group I relate to a method of producing a protein in yeast by recombinant DNA-technique while the "special technical features" of group II relate to an enzyme from *Humicola insolens* which exhibits cellulase activity. These groups of inventions are not so linked as to form a single general inventive concept. There is no technical relationship among those inventions involving one or more of the same or corresponding technical features.
1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
  2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 92/00360

Box III TEXT OF THE ABSTRACT (Continuation of Item 5 of the first sheet)

Method of screening for a DNA sequence coding for a protein of interest, the method comprising

- (a) cloning, in suitable vectors, a DNA library from an organism suspected of producing one or more proteins of interest,
- (b) transforming suitable yeast host cells with said vectors,
- (c) culturing the host cells under suitable conditions to express any protein of interest encoded by a clone in the DNA library, and
- (d) screening for positive clones by determining any activity of a protein expressed in step (c).

An enzyme which exhibits cellulase activity and has been isolated from DNA library of *Hemicola insolens*. The enzyme has a cellulosebinding domain and exhibits endocellulase activity in the presence of linear alkyl benzene sulfonate.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

26/02/93

International application No.  
PCT/DK 92/00360

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s)  | Publication<br>date  |
|---|---------------------|---|--|
| WO-A1- 9201069                            | 23/01/92            | AU-A- 8299191   | 04/02/92   |
| WO-A1- 8909259                            | 05/10/89            | EP-A- 0406314<br>JP-T- 3504080  | 09/01/91<br>12/09/91   |
| WO-A1- 9117243                            | 14/11/91            | AU-A- 7887491   | 27/11/91   |
| US-A- 4435307                             | 06/03/84            | AT-B- 384442<br>AT-B- 394574<br>BE-A- 888632<br>CH-A- 663511<br>DE-A,C- 3117250<br>FR-A,B- 2481712<br>GB-A,B- 2075028<br>JP-C- 1443072<br>JP-A- 57023699<br>JP-B- 61016316<br>NL-A- 8102123 | 10/11/87<br>11/05/92<br>29/10/81<br>31/12/87<br>01/04/82<br>06/11/81<br>11/11/81<br>08/06/88<br>06/02/82<br>30/04/86<br>16/11/81 |